

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平7-327213

(43) 公開日 平成7年(1995)12月12日

(51) Int.Cl.<sup>4</sup>  
H 0 4 N 7/14

識別記号 庁内整理番号

F I

技術表示箇所

審査請求 未請求 請求項の数 3 O L (全 5 頁)

(21) 出願番号 特願平6-117925

(22) 出願日 平成6年(1994)5月31日

(71) 出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72) 発明者 西村 龍志

神奈川県横浜市戸塚区吉田町292番地株式

会社日立製作所映像メディア研究所内

(72) 発明者 織田 真倫子

神奈川県横浜市戸塚区吉田町292番地株式

会社日立製作所映像メディア研究所内

(74) 代理人 弁理士 小川 勝男

(54) 【発明の名称】 テレビ電話装置

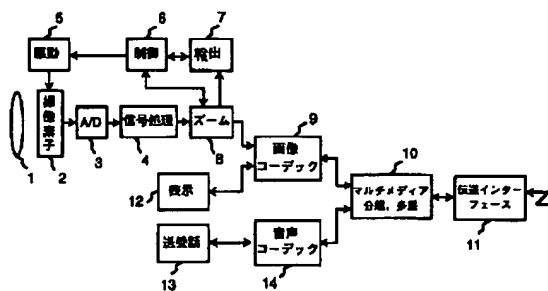
(57) 【要約】

【目的】 通話者の顔画像が再生画面から外れないよう自動制御を行なうテレビ電話を提供すること。

【構成】 テレビ電話装置において、撮像素子の読み出し領域を制御する制御手段と、ビデオ信号を電氣的に拡大するズーム手段と、被写体の位置を検出する検出手段とを設け、上記検出手段の出力する被写体の位置に基づいて撮像素子の読み出し位置を制御するように構成した。

【効果】 曇台等の機械的な手段を用いること無く、電氣的な手段のみで制御を行なうので、通話者の顔画像が再生画面から外れないよう自動制御を行なうテレビ電話を小規模、低消費電力で実現できる。

図 1



## 【特許請求の範囲】

【請求項 1】光学像を電気信号に変換して撮像信号を出力する撮像素子と、  
該撮像素子の出力する撮像信号を用いてビデオ信号を生成する信号処理手段と、  
該信号処理手段の出力するビデオ信号を符号化する符号化手段と、  
該符号化手段の出力する送信ビデオ信号を伝送路へ出力すると共に、伝送路から受信ビデオ信号を入力するためのインターフェース手段と、  
該インターフェース手段から入力された受信ビデオ信号を復号化する復号化手段と、  
該復号化手段の出力するビデオ信号を表示するための表示手段とから構成されたテレビ電話装置において、  
上記撮像素子の特定の領域から撮像信号を読み出し、かつ該特定の領域の位置を変化させるように制御する制御手段と、  
上記撮像素子の特定の領域から読みだされた撮像信号、または該撮像信号から生成されたビデオ信号を電氣的に拡大するズーム手段と、  
被写体の位置を検出する検出手段とを設け、  
上記制御手段は、上記検出手段の出力する被写体の位置に基づいて制御を行なうことを特徴とするテレビ電話装置。

【請求項 2】請求項 1 に記載したテレビ電話装置において、上記検出手段は、被写体の輝度、色情報を用いて被写体位置を検出することを特徴とするテレビ電話装置。

【請求項 3】請求項 1 に記載したテレビ電話装置において、上記検出手段は、被写体の人物の顔を検出することを特徴とするテレビ電話装置。

## 【発明の詳細な説明】

## 【0001】

【産業上の利用分野】本発明はテレビ電話装置に係り、通話者がある程度動いても画角から外れないように自動制御を行なうテレビ電話装置に関する。

## 【0002】

【従来の技術】従来の電話の機能である音声の伝送に加えて映像の伝送も可能にしたテレビ電話、またこれを応用したテレビ会議システムが新しいメディアとして普及しつつある。

【0003】テレビ電話は一般に相手の顔画像だけでなく自分の顔画像もモニタできるように構成されている。しかし通話中は相手の顔画像を見ているため、自分の顔画像が画面の端の方へ動いたり、また極端な場合には画面内から外れてしまっても自分では気付かない場合がある。このような不都合が生じないようにするため、テレビ電話においてカメラ方向を自動的に制御する方式が提案されている。このような技術に関し、例えば特開平 1-300789 号公報に述べられている。

## 【0004】

【発明が解決しようとする課題】しかし、上記従来技術においてはカメラ方向を制御する手段が必要であり、通常このような手段としては電動式の雲台等の機械的な手段が必要となるため、テレビ電話装置の規模が大きくなる問題があった。特に装置の小形化のためカメラをディスプレイの中に組み込むような場合に、カメラ方向を制御できるように構成することは困難であった。

【0005】本発明の目的は、上記問題点を解決し、装置の規模を増大させることなく、通話者が画角から外れないように自動制御を行なうテレビ電話装置を提供することにある。

## 【0006】

【課題を解決するための手段】上記問題点を解決するため本発明では、撮像素子と、該撮像素子の出力する撮像信号を用いてビデオ信号を生成する信号処理手段と、該ビデオ信号を符号化する符号化手段と、該符号化手段の出力する送信ビデオ信号を伝送路へ出力すると共に伝送路から受信ビデオ信号を入力するためのインターフェース手段と、該インターフェース手段から入力された受信ビデオ信号を復号化する復号化手段と、該復号化手段の出力するビデオ信号を表示するための表示手段とから構成されたテレビ電話装置において、上記撮像素子の特定の領域から撮像信号を読み出し、かつ該特定の領域の位置を変化させるようにする制御手段と、上記撮像素子の特定の領域から読みだされた撮像信号、または該撮像信号から生成されたビデオ信号を電氣的に拡大するズーム手段と、被写体の位置を検出する検出手段とを設け、上記制御手段は、上記検出手段の出力する被写体の位置に基づいて制御するようにテレビ電話装置を構成した。

## 【0007】

【作用】上記手段により構成された本発明によるテレビ電話は以下のように動作する。上記撮像素子と、信号処理回路はビデオカメラを構成しており、通話者を撮像してビデオ信号を生成する。生成されたビデオ信号は、符号化手段により符号化された後、インターフェース手段を経て伝送路へ出力され、通話の相手方のテレビ電話装置に送信される。また、受信の場合はインターフェース手段から入力された受信信号を復号化装置で復号化し、表示装置に出力することにより相手方から送られた映像をモニタする。このテレビ電話装置では、撮像された通話者の顔画像が出力画像の画面内から外れてしまうことのないよう、以下のような制御を行なう。

【0008】上記撮像素子は 2 次元撮像素子であり、全体の受光領域よりも小さい特定の領域に蓄積した信号のみを読み出すことが可能となっている。また、この読み出し領域の位置は、上記制御手段により受光領域内で自由に移動できるように構成されている。このように撮像素子が全体の受光領域よりも小さい特定の領域から信号を読み出す場合、ズーム手段は信号処理回路の出力するビデオ信号を電氣的に拡大して再生画像の大きさを一定

にする。また、上記検出手段は、ビデオ信号から被写体である通話者の位置を検出する。制御手段は検出された位置と画面の中心とのずれを計算してこのずれがなくなるよう、すなわち通話者の画像が画面中央にくるように上記撮像素子からの読み出し領域を制御する。以上の動作により、本テレビ電話装置では通話者の顔画像が画面内から外れないように制御することができる。

【0009】

【実施例】本発明の一実施例を図1により説明する。図1は本発明の一実施例を示すテレビ電話装置の構成を示すブロック図であり、1はレンズ、2は撮像素子、3はA/D変換回路、4は信号処理回路、5は駆動回路、6は制御回路、7は被写体検出回路、8はズーム回路であり、これらによって撮像部を構成している。撮像部は、通話者の顔画像を撮像し、CIF(Common Image Format)等、所定のフォーマットに従ったビデオ信号を出力する。

【0010】9は画像コーデックであり、撮像部で生成した送信画像の符号化と受信画像の復号化を行なう。符号化、復号化は、例えばテレビ電話用の符号化標準規格であるCCITTのH. 261等の規格に従って行なう。これらの規格に従って符号化、復号化する方法に関しては、例えばテレビジョン学会誌Vol. 45, No. 7(1991年)第739から799頁、第807頁から812頁に記載されている。また13は送受信器、14は音声コーデックであり、公知の音声信号処理を行なう。音声コーデック14では音声信号の符号化、復号化を行なう。10はマルチメディア分離、多重回路であり、音声信号とビデオ信号について受信の際は分離、送信の際は多重を行なう。伝送インターフェース11は、マルチメディア多重された信号を伝送路へ送受信するためのインターフェースである。

【0011】本発明では、通話者の顔画像が画面から外れないように自動制御するが、この動作について以下に説明する。撮像素子2はCCD(Charge Coupled Device)等の2次元撮像素子であり、レンズ1によってその受光面に結像された光学像を光電変換し、受光面に信号電荷を蓄積する。蓄積された信号電荷は駆動回路5の制御に従って電気的な撮像信号として読み出される。通常、撮像素子2の受光面に蓄積したすべての信号電荷を読み出して、1フィールド(あるいは1フレーム)の画像を生成するが、駆動回路5の制御方法によっては受光面全体よりも小さい特定の領域に蓄積した電荷のみを出力することもできる。こうして読み出した画像を、信号処理において電気的に拡大、補間することにより、電気的なズームが可能となる。このようなズーム処理の方法に関しては例えば1991年テレビジョン学会全国大会予稿集、第359頁から360頁に記載されている。図2により、上記の方法によって通話者の顔画像の位置を制御する方法を説明する。同図(a)は撮像素子2において

撮像された画像の一例を示す模式図であり、このうち、斜線を施していない領域が撮像素子からの読み出し領域である。この読み出し領域の位置は、後述するように被写体検出回路7によって検出された通話者の顔画像の位置に基づいて上下左右に移動させる。こうすることにより、図(b)に示すように、通話者の顔画像が常に画面の中心付近に来るように制御する。

【0012】次に図1における被写体検出回路7について説明する。被写体検出回路7は、入力画像から被写体である人物の顔を検出し、その重心座標等の特徴量を計算して出力する。被写体図3は図1における被写体検出回路7の構成の一例を示すブロック図であり、20~22は2値化回路、26は拡大回路、27は演算回路である。入力端子31からは図2(b)に示す画像に対応した輝度および色信号が入力され、2値化回路20~22において、輝度および色信号の閾値処理を行なって2値化する。2値化回路の閾値は、図1における制御回路6により、インターフェース回路29(図3)を通じて設定される。2値化回路20~22の閾値は、顔画像全体が抽出されるよう、顔面、髪等、輝度と色がほぼ等しいブロックに分けて設定される。その後、2値化された各ブロックをOR回路23で合成することによって、図4(c)に示すような顔全体のシルエットを表わす2値画像が生成される。この2値画像は、AND回路24を経てフィールドメモリ25に1フィールド期間記憶される。フィールドメモリ25に記憶された2値画像は次のフィールドで、拡大回路26により拡大され、AND回路23の一方の入力にフィードバックされる。これにより、検出された被写体が動いても、隣合った2フィールド間で重なり合った領域がある限り、一度検出した被写体を見失うことはない。

【0013】このようにして被写体検出を行ない、その特徴量である重心座標等( $x$ ,  $y$ )を演算回路27によって求める。この結果は、インターフェース回路29を通して、図1における制御回路6に送られる。制御回路6ではこの情報を基に駆動回路5を制御し、撮像素子2の読みだし位置を制御する。このときの制御フローを図5に示す。被写体検出回路7で検出した重心座標等( $x$ ,  $y$ )から、画面の中心座標( $x_0$ ,  $y_0$ )との差分( $\Delta x$ ,  $\Delta y$ )を計算する。この差分を補正するように撮像素子2を制御するには、撮像素子の読みだし開始座標( $x_s$ ,  $y_s$ )を計算し、計算した座標から読み出しが開始されるよう、駆動回路5を制御する。

【0014】本実施例によれば、雲台等の機械的な手段を用いることなく、電気的な手段のみで被写体位置の制御を行なえるので、通話者の顔画像が再生画面から外れないよう自動制御を行なうテレビ電話を小規模、低消費電力で実現できる。

【0015】

【発明の効果】本発明によれば、通話者の顔画像が再生

画面から外れない小規模、低消費電力のテレビ電話を提供できる。

【図面の簡単な説明】

【図 1】本発明による撮像装置の一実施例の構成を示すブロック図である。

【図 2】本発明における撮像画像 (a) と出力画像 (b) の一例を示す模式図である。

【図 3】本発明における被写体検出手段の一構成例を示すブロック図である。

【図 4】被写体検出手段における抽出画像 (a) とその

拡大画像 (b) の一例を示す模式図である。

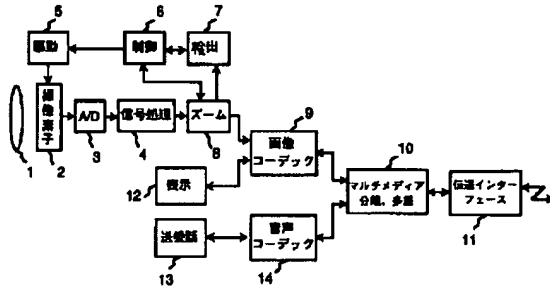
【図 5】本発明における撮像素子の読み出し位置制御の方法を示すフローチャート図である。

【符号の説明】

1…レンズ、2…撮像素子、3…A/D変換回路、4…信号処理回路、5…駆動回路、6…制御回路、7…検出回路、8…ズーム回路、9…画像コーデック、10…マルチメディア分離、多重回路、11…伝送インターフェース、20～22…2値化回路、26…拡大回路、27…演算回路。

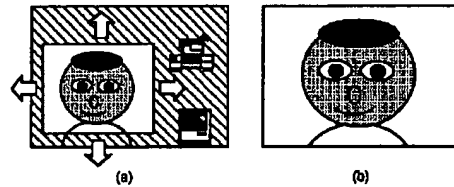
【図 1】

図 1



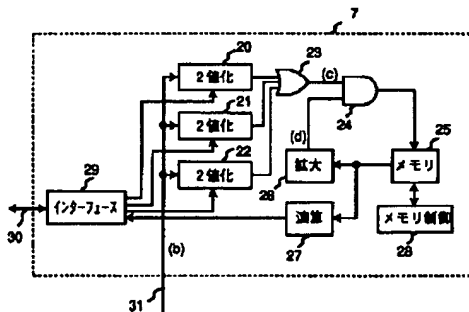
【図 2】

図 2



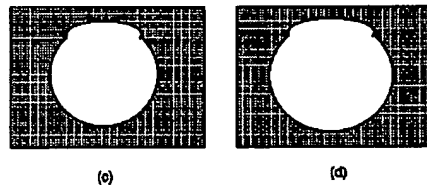
【図 3】

図 3



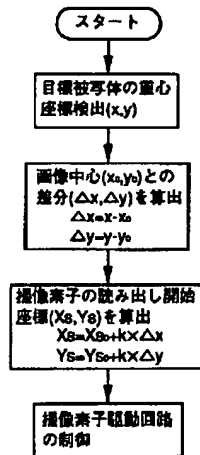
【図 4】

図 4



【図 5】

図 5



撮像素子からの読み出し位置制御フロー

7

Japanese Patent Application Kokai Publication Number

Japanese Patent Laid-Open No. H07-327213

Laid-open date 1995/12/12

Application Number Japanese Patent Application No. H06-117925

Filing date 1994/5/31

Applicant HITAC HI

Inventor: Nishimura, Tatsushi: Oda Mayuko

#### [ABSTRACT]

##### [PURPOSE]

Provide television telephone to automatically control so that face image of a caller is not off reproduction screen.

##### [CONSTITUTION]

In a television telephone equipment, a detecting means to detect zoom lens means to magnify a control means and video signal controlling a reading region of image sensor elements electrically and a position of a subject is established, it was composed a position of the subject which the detecting means output was based on, and to control a reading position of image sensor elements.

##### [Effects]

Because, without using mechanical means such as a universal head, it controls only by electric means, television telephone to automatically control so that face image of a caller does not come from a reproduction screen can be realized in small scale, low power consumption.

##### [Detailed Description of the Invention]

##### [0001]

##### [Industrial Application Field]

A television telephone equipment to automatically control to some extent a person in charge, a caller work to a television telephone equipment as for the present invention, and not to come from angle of view is related to.

##### [0002]

##### [Prior Art]

The television telephone which transmission of a picture enabled as well as the phonic transmission that is a function of a conventional telephone, the video teleconference system which this was applied to again are spreading as the new media.

[0003]

Television telephone is configured so that, in general terms, one's face image can monitor as well as face image of a partner.

However, even if one's face image moves towards an edge of a screen to watch face image of a partner while it speaks over the phone, and it comes from within a screen when it is extreme again, there is the case that is not noticed by oneself. Such an inconvenience does not occur, a method to control a camera direction in television telephone automatically is proposed to need it. For example, about such a technology, it is spoken at Japanese Patent Laid-Open No. 1-300789 bulletin.

[0004]

[Problems to be solved by the Invention]

However, there was the problem that scale of a television telephone equipment grew big so that the means, in the prior art, to control a camera direction was necessary, and mechanical means such as building on high ground of electric was usually necessary for such a means. When a camera is incorporated in a display for miniaturization of a device in particular, what was composed to be able to control a camera direction was difficult.

[0005]

Object of the invention solves the problems, it is to provide a television telephone equipment to automatically control without boosting scale of a device so that a caller does not come from angle of view.

[0006]

[Means to solve the Problems]

An imaging signal begins to be read with the present invention in a television telephone equipment configured from displaying means to display the video signal which decoding means the transmission of a message video signal which coding means to encode signal processing means the imaging signal which image sensor elements and the image sensor elements output is used, and to generate video signal and the video signal and the coding means output is output to transmission channel, and reception video signal input by an interface means to input and the interface means is decoded, and to do reception video signal from transmission channel and the decoding means output from a particular domain of the image sensor elements to solve the problems, a detecting means to detect zoom lens means to magnify generated video signal from and the imaging signal which begins to have been read from a control means changing a regional position of the

specification and a particular domain of the image sensor elements or the imaging signal electrically and a position of a subject is established, a television telephone equipment was composed the control means was based on a position of the subject which the detecting means output, and to control.

[0007]

[Operations]

Television telephone with the present invention which comprised the means works as follows. The image sensor elements and circuit compose video camera, a caller is image-pickuped, and video signal is generated. After the generated video signal was encoded by coding means, pass, and is output an interface means to transmission channel, is transmitted to a television telephone equipment of an opponent of a call. In addition, a decoding does receive signal input in that case of the reception by an interface means in a decoding device, a picture sent to visual display unit by what is output by an opponent is monitored. With this television telephone equipment, carbuncle without face image of an image-pickuped caller coming off from within a screen of an output image, the following control are done.

[0008]

It is image sensor elements, and 2 dimensions of image sensor elements gets possible to read only the signal which accumulated in a small particular domain from a total light receiving domain. In addition, position of this reading domain is configured to be able to move by the control means freely in light receiving domains. As thus described when image sensor elements reads a signal from a small particular domain from a total light receiving region, zoom lens means magnifies the video signal which a circuit outputs electrically, and it is uniformity with size of a reproduction picture. It is similar, and it is done. In addition, the detecting means detects a position of the caller which is a subject from video signal. A reading domain from the image sensor elements is controlled so that a control means calculates a detected position and gap with the center of a screen, and this gap disappears so that in other words a picture of a caller comes to middle of the screen. It can control so that face image of a caller does not come from within a screen with this television telephone equipment by the above-mentioned movement.

[0009]

[Examples]

One embodiment of the invention is explained by FIG. 1. FIG. 1 is a block



diagram to show framing of a tele bidet telephone equipment showing one embodiment of the invention in, and subject detecting circuit, 8 are zoom lens circuits, and, as for 1, as for lens, 2, as for image sensor elements, 3, as for A/D converter circuit, 4, as for circuit, 5, as for drive circuit, 6, control circuit, 7 compose imaging region by these. Imaging department image-pickups face image of a caller, CIF (Common Image Format) outputs video signal in conformity to a predetermined format.

[0010]

9 is picture codec, and encoding of a transmission of a message picture generated in imaging department and a decoding of a reception picture are done. For example, therefore, encoding, decryption are done in a standard such as H.261 of CCITT which is a coding standard standard for television telephone. For example, as for the method that encoding, a decoding do according to these standards, it is described to television society magazine Vol.45, No. 7 (1991) P 739-799, P 807-812. In addition, handset, 14 are voice codec, and 13 does well-known sound signal processing. Encoding of an audio signal, a decoding are done with voice codec 14. 10 is multimedia separation, duplex circuit, and a multiple is done on the occasion of separation, the transmission of a message on the occasion of the reception about an audio signal and video signal. Transmission interface 11 is interface to transmit and receive the signal which multimedia multiplex was able to leave for to transmission channel.

[0011]

According to the present invention, it is automatically controlled so that face image of a caller is not off screen, but, it is equal to or less than it, and this movement is described. As for image sensor elements 2, photo-electric translation does optical image imaged by means of lens 1 which is 2 dimensions of image sensor elements (Charge Coupled Device CCD) by the photoreceptor plane, signal charge is accumulated to a photoreceptor plane. Accumulated signal charge is read as an electric imaging signal according to control of drive circuit 5. All signal charge that accumulated to a photoreceptor plane of image sensor elements 2 is read, and a picture of one field (one or, frame) is usually generated, but, only the electric charge that accumulated in the particular domain that is smaller than the whole photoreceptor plane by a control method of drive circuit 5 can be output. An electric zoom lens is enabled in this way because expansion, an interpolation do a read picture in signal processing electrically. For example, it is described to television Coference Proc. p 359-360 as for a method of such a

zoom processing in 1991. A method to control position of face image of a caller by means of the method by FIG. 2 is explained. The figure above (a) is a schematic block diagram to show an example of a picture image-pickuped in image sensor elements 2, and, of these, the domain where a slanted line is not put for is a reading domain from image sensor elements. Position of this reading domain is based on a position of face image of the caller which, as discussed below, was detected by subject detecting circuit 7, and it is moved in top and bottom right and left. As shown in a figure (b), it controls by doing it this way so that face image of a caller always comes in the vicinity of the center of a screen.

[0012]

Subject detecting circuit 7 in FIG. 1 is explained next. Subject detecting circuit 7 detects a face of the person who is a subject from an input image, quantity of characteristic such as the barycentric coordinate is calculated, and it is output. Subject FIG. 3 is a block diagram to show constructive one example in of subject detecting circuit 7 in FIG. 1, and, the circuit where 20-22 magnify binarization circuit, 26, 27 is an operationed circuit. Brightness corresponding to a picture shown in FIG. 2 (b) and chrominance signal are input from input terminal 31 ; in binarization circuit 20-22, brightness and the threshold of chrominance signal are processed, and binarization is done. The threshold of a binarization circuit is set through interface circuit 29 (FIG. 3) by control circuit 6 in FIG. 1. A face, hair brightness and a color divide it into an approximately equal block, and the threshold of binarization circuit 20-22 is set so that the whole face image is extracted. Binary image expressing a silhouette of the whole face as shown in FIG. 4 is generated a binary afterwards by composing each block considered to be it from OR circuit 23. It is remembered that this binary image can pass with AND circuit 24 to field memory 25 one field period. Binary image memorized to field memory 25 is magnified by enlarged circuit 26 in the next field, is fed back by one input of AND circuit 23. Even if a subject detected by this moves, as far as there is the domain that overlapped between two fields where the neighbor met, a subject detected once needs not to be lost sight of.

[0013]

In this way subject search is done, barycentric coordinate (x, y) which is quantity of the characteristic is demanded with operationed circuit 27. As a result, through interface circuit 29, it is seen off to control circuit 6 in FIG. 1. Drive circuit 5 is controlled in the basis by this information with control circuit 6, of image sensor elements 2, it begins to be read, and a position is controlled. Control flow of this

time is shown in FIG. 5. From barycentric coordinate (x, y) detected with subject detecting circuit 7, a difference ( $\Delta x$ ,  $\Delta y$ ) with a central coordinate (0, x y zero) of a screen is calculated. It is a reading of image sensor elements, and a start coordinate (x s, y s) is calculated to control image sensor elements 2 to compensate this difference, drive circuit 5 is controlled so that reading is started from a calculated coordinate.

[0014]

Because, according to the present embodiment, a subject position is controlled only by electric means without using mechanical means such as a universal head, television telephone to automatically control so that face image of a caller does not come from a reproduction screen can be realized in small scale, low power consumption.

[0015]

[Effects of the Invention]

According to the current invention, the small scale that does not miss face image of a caller from a reproduction screen, television telephone of low power consumption can be provided.

[Brief Description of the Drawings]

[FIG. 1] It is a block diagram showing constitution of one embodiment of imaging device with the present invention.

[FIG. 2] The present invention is a schematic block diagram to show an example of an imaging picture (a) and an output image (b) to be able to put.

[FIG. 3] It is a block diagram showing one constitution example of a subject detecting means in the present invention.

[FIG. 4] It is a schematic block diagram to show an example of an extraction picture (a) in a subject detecting means and the enlarged picture (b) in.

[FIG. 5] It is a figure of Flow chart showing a method of reading position control of image sensor elements in the present invention.

[Denotation of Reference Numerals]

(1) a lens (2), image sensor elements (3), an A/D converter circuit (4), a circuit (5), a drive circuit (6), a control circuit (7), a detecting circuit (8), a zoom lens circuit (9), picture codec (10), multimedia separation, duplex circuit (11), transmission interface (20-22), a binarization circuit (26), an expansion circuit (27), an operationed circuit.